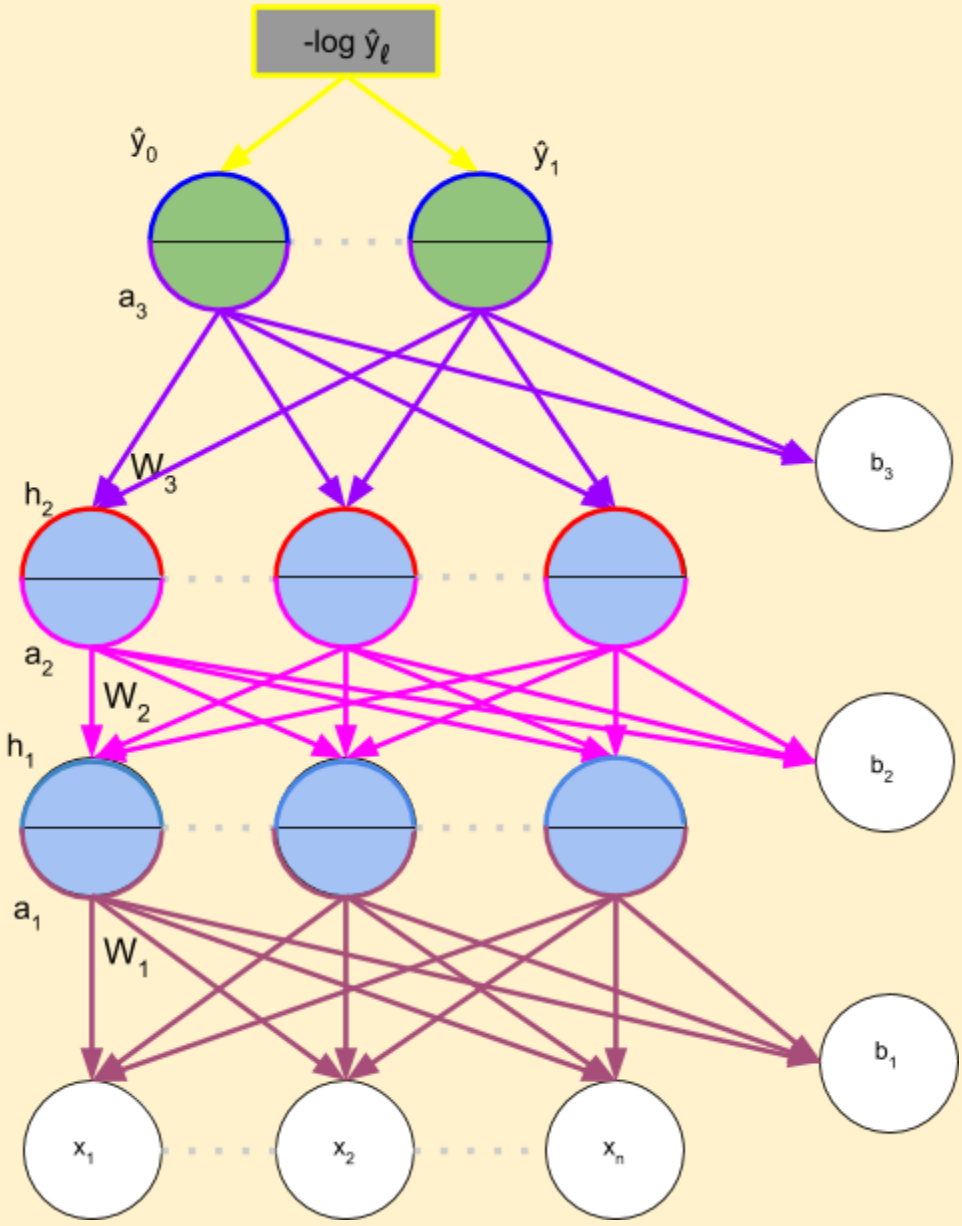
PadhAI: Backpropagation - the full version

One Fourth Labs

A running example of backpropagation

The full story so far

1. Consider the following sample Neural Network



2. Steps to implement Backpropagation
   a. Randomly initialise W and b
   b. Forward propagation
      i. **For k = 1 to L-1 do**
         $a_k = b_k + W_k h_{k-1}$
         $h_k = g(a_k)$
      ii. **end**
      iii. $a_L = b_L + W_L h_{L-1;}$
      iv. $\hat{y} = O(a_L)$

---

c. Backpropagation

i. //Compute output gradient

ii. $\nabla_{a_L} L(\theta) = -(y - \hat{y}_i)$

iii. **For k = L to 1 do**

//Compute gradients w.r.t parameters

$$\nabla_{W_k} L(\theta) = \nabla_{ak} L(\theta) \; . \; h_{k-1}{}^T$$

$$\nabla_{b_k} L(\theta) = \nabla_{ak} L(\theta)$$

//Compute gradients w.r.t layer below

$$\nabla_{h_{k-1}} L(\theta) = (W_k)^T (\nabla_{a_k} L(\theta))$$

/// Compute gradients w.r.t layer below (pre-activation)

$$\nabla_{a_{k-1}} L(\theta) = \nabla_{h_{k-1}} L(\theta) \odot [..., g'(a_{k-1,j}), ...]$$

iv. **end**

3. Sample calculations to be added at a later date. $\Delta y_n / \Delta x_n$